REVISED UG SYLLABUS UNDER CBCS
(Implemented from Academic Year 2020-21)
PROGRAMME: FOUR YEAR B.Sc.,

Domain Subject: **STATISTICS**

Skill Enhancement Courses (SECs) for Semester V, from 2022-23
(Syllabus-Curriculum)

**Structure of SECs for Semester – V**
*(To choose one pair from the three alternate pairs of SECs)*

<table>
<thead>
<tr>
<th>Univ. Code</th>
<th>Courses 6 &amp; 7</th>
<th>Name of Course</th>
<th>Th. Hrs. / Week</th>
<th>IE Mar - ks</th>
<th>EE Mar - ks</th>
<th>Credit s</th>
<th>Prac. Hrs. / Wk</th>
<th>Mar - ks</th>
<th>Credit s</th>
</tr>
</thead>
<tbody>
<tr>
<td>6A</td>
<td></td>
<td>Operations Research - I</td>
<td>3</td>
<td>25</td>
<td>75</td>
<td>3</td>
<td>3</td>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td>7A</td>
<td></td>
<td>Operations Research – II</td>
<td>3</td>
<td>25</td>
<td>75</td>
<td>3</td>
<td>3</td>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td>OR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6B</td>
<td></td>
<td>Statistical Process and Quality Control</td>
<td>3</td>
<td>25</td>
<td>75</td>
<td>3</td>
<td>3</td>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td>7B</td>
<td></td>
<td>Computational Techniques and R Programming</td>
<td>3</td>
<td>25</td>
<td>75</td>
<td>3</td>
<td>3</td>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td>OR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6C</td>
<td></td>
<td>Econometrics</td>
<td>3</td>
<td>25</td>
<td>75</td>
<td>3</td>
<td>3</td>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td>7C</td>
<td></td>
<td>Regression Analysis</td>
<td>3</td>
<td>25</td>
<td>75</td>
<td>3</td>
<td>3</td>
<td>50</td>
<td>2</td>
</tr>
</tbody>
</table>

**Note-1:** For Semester–V, for the domain subject Statistics, any one of the three pairs of SECs shall be chosen courses 6 and 7, i.e., 6A & 7A or 6B & 7B or 6C & 7C. The pair shall not be broken (ABC allotment is random, not on any priority basis).

**Note-2:** One of the main objectives of Skill Enhancement Courses (SEC) is to inculcate field skills related to the domain subject in students. The syllabus of SEC will be partially skill oriented. Hence, teachers shall also impart practical training to students on the field skills embedded in the syllabus citing related real field situations.
Facilities requirements for teaching the Skill Enhancement Courses in Statistics:

All the 6 courses mentioned as Skill Enhancement Courses require a computer lab with at least 20 desk top systems since each student has to practice the methodology and acquire skills in producing output. This requires various software packages both standard (like Microsoft Excel, TORA/ LINDO/LINGO for Operations Research,) as well as open-source packages (like R and Python) to handle statistical analysis. Once these facilities are provided and training given on these lines, the outgoing students will fit into skill oriented jobs like business data analysts and Data Scientists.
Objective: The Objective of the paper is to introduce the basic concepts of Operational Research and linear programming to the students.

Learning Outcomes:
After learning this course, the student will be able
1. To know the scope of Operations Research
2. To link the OR techniques with business environment and life sciences
3. To convert real life problems into mathematical models
4. To find a solution to the problem in different cases
5. To inculcate logical thinking to find a solution to the problem

UNIT-I

UNIT-II
General linear programming Problem(GLP) – Definition and Matrix form of GLP problem, Slack variable, Surplus variable, unrestricted Variable, Standard form of LPP and Canonical form of LPP. Definitions of Solution, Basic Solution, Degenerate Solution, Basic feasible Solution and Optimum Basic Feasible Solution. Introduction to Simplex method and Computational procedure of simplex algorithm. Solving LPP by Simplex method (Maximization case and Minimization case)

UNIT-III
Artificial variable technique - Big-M method and Two-phase simplex method, Degeneracy in LPP and method to resolve degeneracy. Alternative solution, Unbounded solution, Non existing feasible solution and Solution of simultaneous equations by Simplex method.
UNIT-IV

Duality in Linear Programming –Concept of duality -Definition of Primal and Dual Problems, General rules for converting any primal into its Dual, Economic interpretation of duality, Relation between the solution of Primal and Dual problem(statements only). Using duality to solve primal problem. Dual Simplex Method.

UNIT-V

Post Optimal Analysis- Changes in cost Vector C, Changes in the Requirement Vector b and changes in the Coefficient Matrix A. Structural Changes in a LPP.

Reference Books:
5. HadlyG :Linear programming. Addison-Wesley.

Practical/Lab to be performed on a computer using OR/Statistical packages

1. To solve Linear Programming Problem using Graphical Method with
   (i) Unbounded solution
   (ii) Infeasible solution
   (iii) Alternative or multiple solutions.
2. Solution of LPP with simplex method.
4. Problem solving using Two Phase method.
5. Illustration of following special cases in LPP using Simplex method
   (i) Unrestricted variables
   (ii) Unbounded solution
   (iii) Infeasible solution
   (iv) Alternative or multiple solutions.
7. Problems based on Dual simplex method.
Objective: To enrich the knowledge of students with advanced techniques of linear programming problem along with real life applications.

Learning Outcomes:

After learning this course, the student will be able

1. To solve the problems in logistics
2. To find a solution for the problems having space constraints
3. To minimize the total elapsed time in an industry by efficient allocation of jobs to the suitable persons.
4. To find a solution for an adequate usage of human resources
5. To find the most plausible solutions in industries and agriculture when a random environment exists.

UNIT -I


UNIT-II


UNIT-III

UNIT-IV

Network Scheduling: Basic Components of a network, nodes and arcs, events and activities– Rules of Network construction – Time calculations in networks - Critical Path method (CPM) and PERT.

UNIT –V


Reference Books:
7. Dr.NVS Raju; Operations Research, SMS education,

Practical/Lab to be performed on a computer using OR/Statistical packages

1. IBFS of transportation problem by using North- West corner rule, Matrix minimum method and VAM
2. Optimum solution to balanced and unbalanced transportation problems by MODI method (both maximization and minimization cases)
3. Solution of Assignment problem using Hungarian method (both maximization and minimization cases),
4. Solution of sequencing problem—processing of n jobs through two machines
5. Solution of sequencing problem- processing of n jobs through three machines
6. To perform Project scheduling of a given project (Deterministic case-CPM).
7. To perform Project scheduling of a given project (Probabilistic case-PERT).
8. Graphical method of solving for m x 2 and 2 x n games.
9. Solution of m x n games by dominance rule.
10. Solution of n x n games by using matrix oddment method.
11. Linear programming method for solving m x n games.
Course Objectives: To understand the concept of quality, process control and product control using control chart techniques and sampling inspection plan. To have an idea about quality management, quality circles, quality movement and standardizations for quality.

Learning Outcomes:
After learning this course, the student will be able
1. To define ‘quality’ in a scientific way
2. To differentiate between process control and product control
3. To speak about quality awareness in industry
4. To pave a path to an industry to meet the standards
5. To effectively implement various plans to control the quality standards at various stages of an industry.

Unit I
Meaning of quality, concept of total quality management (TQM) and six-sigma, ISO, comparison between TQM and Six Sigma, Meaning and purpose of Statistical Quality Control (SQC), Seven Process Control Tools of Statistical Quality Control (SQC) (i) Histogram (ii) Check Sheet, (iii) Pareto Diagram (iv) Cause and effect diagram (CED), (v) Defect concentration diagram (vi) Scatter Diagram (vii) Control chart. (Only introduction of 7 tools is expected).

Unit II
Statistical basis of Shewhart control charts, use of control charts. Interpretation of control charts, Control limits, Natural tolerance limits and specification limits. Chance causes and assignable causes of variation, justification for the use of 3-sigma limits for normal distribution, Criteria for detecting lack of control situations:
(i) At least one point outside the control limits
(ii) A run of seven or more points above or below central line.

Unit III
Control charts for Variables: Introduction and Construction of $\bar{X}$ and R chart and Standard Deviation Chart when standards are specified and unspecified, corrective action if the process is out of statistical control.
Control charts for Attributes: Introduction and Construction of p chart, np chart, C Chart and U charts when standards are specified and unspecified, corrective action if the process is out of statistical control.
Unit IV
Acceptance Sampling for Attributes: Introduction, Concept of sampling inspection plan, Comparison between 100% inspection and sampling inspection. Procedures of acceptance sampling with rectification, Single sampling plan and double sampling plan. Producer's risk and Consumer's risk, Operating characteristic (OC) curve, Acceptable Quality Level (AQL), Lot Tolerance Fraction Defective (LTFD) and Lot Tolerance Percent Defective (LTPD), Average Outgoing Quality (AOQ) and Average Outgoing Quality Limit (AOQL), AOQ curve, Average Sample Number (ASN), Average Total Inspection (ATI).

Unit V

Text Books:

References:
1. R.C. Gupta: Statistical Quality Control.

Practical/Lab to be performed on a computer using Statistical packages
1. Construction of $\bar{X}$ and R Charts.
2. Construction of $\bar{X}$ and $\sigma$ Charts.
3. Construction of p Charts for fixed sample size.
5. Construction of np Charts.
7. Construction of U charts.
8. Single sampling plan for attributes (OC Curve, Producer’s and Consumer’s risks, AOQ, AOQL, ATI).
9. Determination of single sampling plan by: a) lot quality approach b) average quality approach.
10. Double sampling plan for attributes (OC curve, AOQ, AOQL, ATI, ASN using Poisson distribution).
Course Code: Four-year B.A./B.Sc.
Domain Subject: Statistics
Course 7B: Computational Techniques and R Programming
(Skill Enhancement Course (Elective), 05 Credits
Max. Marks: Theory :100 + Practicals: 50

Course Objectives: To learn the statistical analysis with the help of the statistical software R

Learning Outcomes:
After learning this course the student will be able
1. Understand the basic functioning of a computer
2. Acquire skills in handling business and organizational data using Excel
3. Perform simple analytics using Excel
4. Understand the power of R programming language
5. Handle various statistical issues using R language

Unit I
Computer basics: Basic applications of computer, components of computer system, Central Processing Unit (CPU), input and output units, computer memory and mass storage devices. Programming languages and their applications. Concept of files and folders. Software and types of software. Operating System like Windows and Linux.

Unit II
Data processing using spreadsheets – Data entry and editing features in Excel, copy, paste, paste special options, sort and filter options, auto sum, steps of finding average and standard deviation of data using statistical functions. Matrix operations like transpose, multiply and inverse using Excel functions. Simple graphs like bar chart, line chart and pie chart in Excel. Exporting Excel output to word processors like MS-Word and slide presentations like Power Point.

Unit III
Scatter diagram, fitting of straight line, polynomial and power curves using Excel – Reading R-square value and equation from the graph. Predicting future values using ‘forecast’ and ‘trend’ functions. Data Analysis Pak and its features. Performing Student’s t-test and one-way Analysis of Variance using Data Analysis Pak. P-value and its interpretation.

Unit IV
Programming with R: Introduction to R, Data types in R (numeric, logical, character, complex etc.), R objects: vector, matrix, array, list, data frame, factor, and time series. Arithmetic, logical and relational operators, explicit and implicit looping, functions and functional programming in R, Lexical scoping rules in R, benefits of Lexical scoping, other scoping rules, debugging facility in R. Few important mathematical, statistical and graphical functions in R.
Unit V
Descriptive Statistics with R software: Calculations with R software such as descriptive statistics, frequency distribution, Graphics and plots, statistical functions of central tendency, variation, skewness and kurtosis and illustration with examples.

Suggested Books


Practical/Lab to be performed on a computer using R Software

1. Construction of Bar Chart and Pie Chart using Excel
2. Fitting of straight-line using Excel
3. Calculating Matrix Inverse using Excel
4. One way ANOVA using Excel
5. Data visualization using R - frequency polygon, Ogives, Histogram.
6. Data visualization using R - simple and multiple bar diagram, pie chart.
8. Computation of Karl Pearson’s Coefficient of Correlation and Rank Correlation using R.
9. Construction of Control Charts for variables ($\bar{X}$, R and $\sigma$) charts using R.
10. Construction of Control Charts for attributes (p, np charts with fixed and varying sample sizes) using R.
Learning Outcomes:

The course on econometrics will primarily focus on the use of statistical modeling and the relevant analyses to economic data problems. After learning this course the student will be able understand

1. various important econometric models and relevant model building concepts in econometrics
2. general linear models and estimation of inherent model parameters
3. multicollinearity, its detection and consequences and related inferential aspects
4. some advanced concepts of generalised least squares estimation, autocorrelation, its consequences, detection and strategy for reducing autocorrelation,
5. heteroscedasticity and its inherent concepts including its consequences,
6. some inferential aspects on heteroscedasticity,
7. practical aspects and real data illustration of the related problems.

UNIT-I

Basic Econometrics: Nature of econometrics and economic data, concept of econometrics, steps in empirical economic analysis, econometric model, importance of measurement in economics, the structure of econometric data, cross section, pooled cross section, time series and paired data.

UNIT-II


UNIT- III


UNIT- IV

**Multicollinearity**: The concept of multicollinearity and its consequences on econometric models, detection of multicollinearity. Measure of Multicollinearity – Variance Inflation Factor (VIF) and tolerance, formula and interpretation. Methods of reducing the influence of multicollinearity.
UNIT - V

**Autocorrelation:** Disturbance term (u) in econometric models and its assumptions, autocorrelated disturbances and their consequences on the model parameters, Detecting the presence of autocorrelation – hypothesis tests for autocorrelation - Durbin Watson test and its interpretation.

**References:**


**Practical/Lab to be performed on a computer using Statistical packages**

1. Problems based on estimation of General linear model.
2. Testing of parameters of General linear model.
3. Forecasting of General linear model.
4. Problems concerning specification errors.
5. Problems related to consequences of Multicollinearity.
6. Diagnostics of Multicollinearity.
7. Problems related to consequences of Autocorrelation (AR(I)).
8. Diagnostics of Autocorrelation.
11. Diagnostics of Heteroscedasticity.
Learning Outcomes:

After learning this course the student will be able

1. To know about regression techniques, which are powerful tools in statistics,
2. To get an idea of Linear and Multiple Linear regression,
3. To learn about regression diagnostics, residual plots for visualization
4. To perform statistical tests of hypotheses on regression coefficients.
5. To study the structural stability of a regression model.
6. To learn the regression with qualitative independent and dependent variables by dummy variable technique.
7. To learn the selection of the best regression model.

Unit I
Simple Linear Regression: Simple Linear Regression Model. Least-Squares Estimation of the Parameters - Estimation of $\beta_0$ and $\beta_1$, Properties of the Least-Squares Estimators and the Fitted Regression Model. Hypothesis Testing on the Slope and Intercept - Use of t Tests, Testing Significance of Regression and Analysis of Variance

Unit II

Unit III
Regressions with Qualitative Independent Variables: Use of dummy variables to handle categorical independent variables in regression. Estimation of model parameters with dummy variables - Testing the structural stability of regression models, comparing the slopes of two regression models. Multiple linear regression with interaction effects.

Unit-IV
Regressions with Qualitative Dependent Variables: Binary logistic regression with several independent variables, estimation of coefficients, evaluating the Odds Ratio (OR) and its interpretation. The concept of Piecewise linear regression, The Logit, Probit and Tobit models and their applications.
Unit – V

Best Model: Selecting ‘Best’ regression model. All possible regressions – $R^2$, Adjusted $R^2$, $MS_{Res}$, Mallow’s statistic. Sequential selection of variables – criteria for including and eliminating a variable – forward selection, backward elimination and stepwise regression.

References:


Practical/Lab to be performed on a computer using Statistical packages

1. Least Squares estimates of slope and intercept
2. Plotting of two Regression Lines
3. Finding $R$-square value of Linear Models
4. Student’s t-test for regression coefficient
5. ANOVA for Multiple Linear Regression model
6. Selecting best regression model by $R^2$
7. Selecting best regression model by Adjusted $R^2$
8. Selecting best regression model by $MS_{Res}$
9. Selecting best regression model by Mallow’s statistic
10. Selecting best regression model by forward selection
11. Selecting best regression model by backward elimination.